

## **The Political Morality of Nudges in Healthcare<sup>1</sup>**

Jonathan Gingerich

May 25, 2015

### **I. Introduction**

Nudges have sparked the interest of some moral philosophers, many of whom have thought that there is something ethically unsettling about nudges because of the sort of relationship between a government and its citizens that they entail or because of what nudges do to the deliberative capacities of people who are nudged. Daniel M. Hausman and Brynn Welch argue that nudges “may threaten the individual’s control over her own choosing” and claim that to “the extent that they are attempts to undermine that individual’s control over her own deliberation, as well as her ability to assess for herself her alternatives, they are *prima facie* as threatening to liberty, broadly understood, as is overt coercion” (2010). Jeremy Waldron suggests that there is a “genuine worry” about Cass Sunstein’s advocacy of nudges, and thinks that there is “an element of insult” in Sunstein’s nudging (2014). For these and other philosophers, nudges raise ethical concerns primarily or exclusively when used by governments to encourage their citizens to behave in particular ways and governments have at least a *pro tanto* reason to avoid nudging their citizens.

In this paper, I reconstruct the moral philosophers’ objections to nudges, explaining why we might think that nudges contain “an element of insult.” I begin by considering and rejecting as overly simplistic interpretations of the moral philosophers’ objection that focus on either how nudges diminish the number of choices that people make or how nudges aim to modify people’s

---

<sup>1</sup> Thanks to Glenn Cohen, Daniela Dover, Rimon Elkotbeid, Tatiana Espinosa, Nir Eyal, Laura Gillespie, Pamela Hieronymi, Melissa Hughs, Brian Hutler, Christopher Robertson, Yashar Saghai, Seana Shiffrin, and the audience at the Petrie-Flom Center’s 2014 Annual Conference.

behavior without engaging with their rational capacities. I suggest that a more charitable interpretation of the moral philosopher's objection, according to which nudges are morally worrisome insofar as they undermine democratic control of default rules by making it more difficult for people to detect that the choice architecture in which they operate has been actively shaped by policymakers. I argue that this version of the objection to nudges is not really a worry about nudges, as a policy category, but a worry about compromising the integrity of individual decision-making in certain important contexts. Governments do not, then, have a *pro tanto* reason to refrain from nudging. Instead, governments ought to consider whether particular nudges prevent people from directly engaging with the reasons and values that bear on a decision that it is morally important for them to make for themselves. Finally, I discuss how governments might determine whether nudges in the field of healthcare are likely to interfere with the independence of such decisions.

## **II. Nudges and Respect for Choice**

A central feature of moral philosophers' criticisms of nudges is a claim that, in some manner, nudges fail to respect nudgees as individual choosers, whether by failing to show appropriate regard for the choices that they make or by treating people as lacking the rational capacities that are central to the practice of choosing.<sup>2</sup> Waldron, for instance, claims that when I am nudged, "my choosing is being made a mere means to my ends by somebody else" (Waldron 2014). Waldron "think[s] this is what the concern about dignity is all about" (Ibid.). My choosing is treated as a mere means when a nudge fails to take seriously my capacities for deliberation as a part of me, and the fact that when I do something because I choose to do it, I invest my action with self-respect (Ibid.). In this section, I explore why we might think that

---

<sup>2</sup> I use the terms "nudger" to designate a person or institution who institutes a nudge and "nudgee" for a person who is the target of a nudge.

nudges fail to respect my choosing. I first consider two obvious interpretations of the moral philosophers' objection to nudges and reject both as overly simplistic. I then propose that the best understanding of the objection to nudges focuses on the possibility that nudges could undermine democratic control of choice architecture.

*A. What the Moral Philosophers' Objection to Nudges Is Not*

One very straightforward interpretation of the objection to nudges, which is advanced by some libertarian critics of nudging, holds that nudges disrespect me as a chooser because they reduce the number of options that I have to choose from, or the number of choices that I must make. But it is a truism that having more choices does not straightforwardly enhance my agency. In many contexts, having more choices limits both my welfare, by forcing me to make choices I would rather not make (like which health insurance plan to enroll in), and my autonomy, by forcing me to expend deliberative resources making choices that are not central to my agency (Sunstein 2015). If I am nudged by my insurance to choose an optometrist from a relatively small network of optometrists, I might spend less time deliberating about which dentist to use and more time deliberating about things that matter much more to me. So the complaint that nudges fail to respect nudgees as choosers cannot plausibly amount to the claim that nudges reduce the number of choices that people have.

We might instead think that this objection concerns the manner that nudges interact with individuals' rational choice-making capacities. Nudges do not induce nudgees to behave in manner that promotes their own or third-party welfare or deliberation through rational persuasion. Sometimes nudges consist in providing information, like graphic warnings on cigarette packages about the health effects of smoking. But nudges differ from argumentation in that, even when they consist in the providing information, they achieve their objectives through sub-rational

processes. Graphic cigarette warnings “nudges” smokers to pay more attention to the long-term health effects of smoking not by providing an explanation of why they should focus my reasoning about cigarette purchases on long-term health consequences, but by relying on my reflexive response to the warning. Nudges may constitute a substitution of nudger’s judgment for my own, such that “my choosing is being made a mere means to my ends by somebody else” (Waldron 2014), because they aim to elicit a specific behavior from me other than by rationally persuading me that I should behave in that way.

While this interpretation of the objection to nudges is more plausible than the first version that I considered, it remains unappealing. Our human rational agency is inevitably conditioned on and influenced by non-rational processes. Perhaps we should aspire to free our rational decisions from non-rational influences to the greatest extent possible, but it would be odd to think that the best way to do this is to reduce the effect of extraneous influences on every single choice we make. Further, it is not clear why we would want to avoid all non-rational interferences with deliberation that aim at a particular result. As the proponents of nudging frequently point out, even in the context of important decisions that should be made independently, some default rule is often necessary. Deciding whether to become an organ donor is a decision that many people think is an important one that individuals ought to make autonomously and with the use of their own rational, deliberative capacities. But what should happen with the organs of someone who has not made a decision about whether to be an organ donor or not? Given that having some default rule is inevitable, should the default be that people are organ donors or that they are not? And if, as humans, it is inevitable that non-rational processes impact individual’s rational decision-making so that many people stick with the default rule for organ donation, whatever it is, what rationale could there be for setting the default at not

being a donor if being a donor is what most people would prefer if they did, in fact, give serious consideration to the question of whether to be an organ donor? It seems that even the fact that a nudge has the potential to make a particular outcome to a particular *important* decision more likely (other than by rational persuasion) is not, by itself, reason to be suspicious of nudges.<sup>3</sup>

### *B. Nudges and Democratic Control*

Part of why nudges seem unlikely to impermissibly interfere with the independence of reasoning by individuals any more than rational persuasion or mandates is that nudges tend to succeed and fail at the level of populations. Consider an employer that changes the health insurance plan that its employees are enrolled in by default from Plan A, a low premium, high deductible plan, to Plan B, a moderate premium, moderate deductible plan, although it allows employees to easily choose which of the two plans to enroll in. The employer makes the change because such a plan is likely to be better, financially speaking, for most of its employees. Whether changing the default insurance plan succeeds or fails does not depend on whether employee X sticks with the default or elects to enroll in a different plan, and it also does not depend on whether employee X would be better off with Plan A or Plan B. There would be no good ground to object to the nudge if, in fact, most employees would like to enroll in Plan B and only a small number of employees would like Plan A, and if setting Plan A as the default will result in more employees enrolling in Plan A than Plan B. The employer could know that most of its employees would prefer Plan B--for instance, by surveying a random sampling of employees--it is impossible for the employer to know the health insurance preferences of all of

---

<sup>3</sup> I leave open the possibility that in some contexts, such as decisions about organ donation or other posthumous uses of one's body might involve an asymmetry, such that a default rule of allowing posthumous use of a body is intrusive even if, upon considered reflection, almost everyone would not have an objection to using their organs after their death. Thanks to Glenn Cohen for raising this point.

its employees in advance. The default is going to be the insurance plan that some employees want but not what others want, and there must be some default. Setting Plan B as the default looks like the best way to allocate the costs of decision-making to reduce costs as much as possible.

Focusing on how nudges function at the level of populations suggests that if nudges give rise to a moral concern, the concern has to do not with their impact on individual decision-making but on how they impact the choice architecture that a group of people relies on. Defaulting to Plan A seems like a good rule to have if a plurality of employees would, if they thought about it, save at that rate. But consider a nudge to encourage employees to avoid financial shocks associated with explained healthcare expenditures by setting Plan B as the default when, in fact, most employees would choose Plan A if they really thought about it. This nudge looks much more like it fails to respect individual members of the population of employees as choosers. The objection to nudging could be formulated as:

N1: Nudges disrespect nudgees as choosers when they make it more difficult for individual nudgees to choose according to the default choice that the collective of nudgees would prefer.

Here, “the default choice that the collective of nudgees would prefer” represents the appropriate make based on the appropriate aggregation of individual nudgees’ choices. If a doctor routinely prescribes two versions of a drug, generic and branded, and 90% of patients want a generic and 10% want a brand-name drug, the majoritarian default choice would be the generic. According to this line of thought, nudges disrespect us as choosers when they treat our choices as something

to be modified in pursuit of some further objective that we would not endorse, individually or together.

This objection needs further refinement because often there may be no fact of the matter as to what we would endorse, individually or together. The problem is not only that, given certain assumptions about how to aggregate individual preferences, it may be impossible to translate individual preferences into community-wide preferences, but also that individuals may not have an answer as to what their highest order desires are (Korobkin 2009). An insurance plan may serve many patients who both want to want the best medicine available (and who think brand-name drugs are better than generics) and who want to want good healthcare at an affordable price (and who think this is provided by generics). It may be a rational failing of individuals to have unresolved conflicts between their preferences, but one can purchase pharmaceuticals without being fully rational. Even if they were prompted to reflect seriously about their higher order preferences, many customers would not know what they prefer, or might simply pick among alternatives for no reason at all. More abstractly, I might not have fully consistent desires about what choices I would like to make. For instance, I might like to make more choices than I possibly could make, given the limited amount of time and the limited deliberative resources that I have, and I might not have worked out rules about how to prioritize the various choices that I would like to make. Additionally, figuring out which choices are the most important for me to make is the work of a lifetime, not something that I could settle by setting aside a couple of hours or even a few days to think about it. Moreover, in many cases, there might not be any default choice that represents an aggregation of individual preferences about the default, even in principle. If this is the case, we might accept that N1 is only an objection to nudges in those cases where there is a clear majoritarian default choice. But

thinking about the grounds of a principle that supports showing respect for majoritarian default choices can produce a revised objection. Nudges can fail to respect people as choosers when they promote some small set of individuals' default preferences at the cost of making it (at least marginally) more difficult for other people to get what they want. What matters is democratic control of nudges so that it is possible for us, together, to revise nudges if we want to do so. The objection to nudges might then be formulated as:

N2: Nudges disrespect nudgees as choosers when they make it more difficult for nudgees to democratically control the relevant choice architecture.

Nudging patients toward generics and away from name-brands would, then, be subject to criticism under N2 if most patients wanted branded drugs, or if most patients were not sure what they wanted to want but the nudge made it more difficult for them to exercise control over decisions about the default drugs in prescriptions. Of course, there are many different ways that a nudge could be under democratic control: a nudge might be implemented through a referendum, through a statute enacted by a legislature, or by an administrative agency overseen by an elected executive. There are many different forms that democratic oversight can take, but more significant to determining whether a nudge is subject to criticism under N2 is the aim of the nudge. N2 objects to nudges that have an anti-democratic aim in that they treat some people's preferences or commitments as more important than those of others in determining a default, just because of who they are.

At this point, we might wonder why nudges, in particular, are a form of policy-making that we should worry about. Why think that there is a *pro tanto* reason to avoid nudging?

Perhaps the thing to avoid is not *nudges*, but rather failures of democratic oversight of policymaking, or counter-democratic aims of government officials.

For the moral philosophers' objection to nudges to work as an objection to nudges particularly, it should be able to explain why there is something more morally worrisome about nudges than about coercive policies. The focus on democratic control of choice architecture helps to explain how this might happen. As non-lying deception may be more difficult to detect than lies, it may be harder for nudges to tell when nudging is going on than to tell when more coercive policies are in place.<sup>4</sup> If a hospital is designed so that its stairs and ramps are attractive and obvious, so that patients, visitors, and staff are more likely to walk than take an elevator or lift, it is hard for users of the hospital to immediately discern that the placement of stairs is the result of a policy that has an aim related to how they should get exercise. Even if they can tell, in the abstract, that the placement of the stairs represents a policy decision, the policy intervention is much less salient than would be a prohibition on using elevators by anyone who could take the stairs, or, for that matter, a requirement that elevator passengers fill out a waiver acknowledging the health risks of not walking enough before going up a few stories. If the presence of nudges in our deliberative environments is more difficult to detect than the presence of more overtly coercive policies or than the presence of efforts at rational persuasion, we might think that nudges can undermine democratic control of choice architecture. Rather than thinking about the objectives that are being advanced by making stairs obvious and attractive and doing something to change the hospital's architectural decisions if they disagree with its objectives, hospital users are likely to go along with the nudge either because they do not notice that it is part of the choice architecture or, even if they notice it, because individually, they still have the freedom to take the

---

<sup>4</sup> I owe this analogy to Seana Shiffrin.

elevator and so are not very concerned to change the policy. The involvement of my choice in the outcome may make it less likely that I will object to the manner that the choice was set up for me. After all, in the end I got what I wanted and at relatively little cost to me.

We might wonder whether nudges rely on this quality of non-transparency to function. In many cases, greater transparency might undermine nudges that are *prima facie* objectionable according to N2. If, collectively, the people enrolled in the insurance plan that nudges them to use generics do not want to use generics, the disclosure of the nudge might prompt a large number of insureds to actively choose name-brand drugs. There is a complication, however: the members of the insurance plan might prefer brand-name drugs but also have a higher-order desire not to have this preference. They might think that they themselves are irrational for wanting branded drugs when generics produce health outcomes that are just as good, and so, collectively, they might want not only a generic drug default but also want not to be prompted to think about the possibility of opting out by a disclosure. We might also think that nudges do not depend on being hidden to work because even when disclosure about a nudge is provided, nudges are under pressure to see themselves as responsible for what they choose. Suppose I work for an employer that offers health insurance through Plan A (the low premium, high deductible plan) and Plan B (the moderate deductible, moderate premium plan), and makes it extremely easy to select either plan, but offers Plan A as the default. If I would rather have Plan B but never fill out the very simple form to enroll in Plan B, whether out of procrastination, anxiety, laziness, or inattentiveness, and I end up with high out of pocket medical expenses that I would have avoided if I selected Plan B, I might blame my employer for not making Plan B the default, but I might also blame myself. Disclaiming responsibility would involve undermining

my own rational agency and foregrounding my rational defects, a painful experience that I might seek to avoid.

We can now formulate a moral claim that is more specifically focused on nudges:

N3: Nudges disrespect nudgees as choosers when they make it more difficult for nudgees to democratically control the relevant choice architecture by making the manipulation of majoritarian default rules less obvious to nudgees.

This principle echoes Nir Eyal's suggestion that nudges are morally fine when they put nudgees on "automatic pilot" toward a particular course of action but make it easily possible for nudgees to abort automatic pilot (2015). N3 suggests that nudges are morally objectionable when they make it difficult for us, collectively, to get off of autopilot. To the extent that nudges make us feel all right about being governed from above, and make it less likely that we will organize democratically to change policies that do not suit us, we should be concerned about nudges.

### **III. Nudges and Healthcare**

Some difficulties with the objection to nudging linger. Suppose that a city with many employees offers five health insurance plans, A, B, C, D, and E. The city changes the default insurance plan from A to B, because public health experts tell the city that this will enhance the welfare of its employees in the long run. And suppose that this is a case where there is no default rule that the community of nudgees would or could democratically endorse--most employees do not know which plan they would choose if they thought about it, and even if they spent a lot of time deliberating about which plan to choose, most employees would end up just picking a plan more or less at random. Setting B as the default plan might reduce democratic

control of the choice architecture relative to a decision to force employees to make a choice among the five plans as a condition of employment or relative to a penalty default rule, like making the default insurance plan the one that is obviously worse than the others. It seems strange to think that the city's nudge is morally problematic. It is strange to think that it is important to maintain a high level of democratic control over every single piece of choice architecture, at least as it remains, in principle, possible to exert democratic control. If anything troublesome were happening in the case of the city's health care plans, it would be because the choice that the nudge is designed to shape is one that it is particularly important for individuals to make for themselves, by directly engaging with the values that bear on their decision. Seana Shiffrin suggests that, "It is valuable to have the opportunity to engage with a particular value, in some degree of isolation, to determine its significance to oneself and to respond appropriately to the reasons it presents" (Shiffrin 2004). On this picture, choice architecture should not be designed in a manner that makes it difficult for people to engage directly with the reasons and values that bear on a decision that they face, when the area of decision is one where it is morally important for individuals to practice such direct engagement.

Consider a cancer patient who has to choose between two courses of treatment, one palliative and one involving aggressive chemotherapy, both of which are routinely provided to other patients with the same condition. It might be important that the patient's deliberations about which course of treatment to pursue without being pressured by considerations about whether her family will be bankrupted or severely financially strained by one or another of the treatment options, or about whether others will disapprove of or be inconvenienced by her decision.<sup>5</sup> We might think it important instead for the patient to focus directly on values related

---

<sup>5</sup> This example closely tracks and example provided in Shiffrin 2004, 289.

to her own health and longevity, and some nudges might interfere with the patient's ability to do so. Indeed, in this situation, providing the patient with either course of treatment as a default option might signal to the patient that her healthcare providers would disapprove of her decision to pursue the other course of treatment.

Sometimes, when things are very important to us, we are unlikely to be swayed by a default rule, especially one that provides a very easy mechanism for opting out, because a decision that is very important to us is precisely the sort of thing that we are likely to actively deliberate about. However, there are two reasons that we might remain concerned about the affect of nudges on decision-making. First, even in decisions that I regard as very important, nudges might interfere with my direct engagement with values, even if they do not prevent deliberation. Setting the default rule for the aforementioned cancer patient as providing aggressive chemotherapy might not make it much less likely that the patient will actively deliberate about which course of treatment to pursue, but it might undermine the patients ability to engage in a focused deliberation about the values of health. She might take the setting of the default rule to indicate that it is appropriate for patients to want aggressive chemotherapy, or that her physicians endorse of this treatment option.

Second, some decisions that are really important to me can also be very emotionally fraught, especially in the context of healthcare. Deciding between aggressive chemotherapy and palliative treatment forces me to confront many different values that may be incompatible with one another. I might care about having a chance to prolong my life, about the environment in which I receive healthcare, about avoiding severe pain, and about maintaining cognitive function, and I may not be able to preserve access to all of these values that I care deeply about. This could be profoundly anxiety provoking, and might occasion avoidance. If a default is available, I

might take the default as authoritative so that I need not endure the stress of directly engaging with the values that bear on a decision about my course of treatment. This does not end the inquiry about whether such a nudge is desirable, because we might have further values that are served by making it easy for patients to avoid the anxiety provoked by making emotionally difficult decisions about medical treatment. But it is possible that, while collectively we recognize that when faced with such decisions we would be likely to avoid them by relying on defaults available to us, we might also collectively think it important that we overcome this avoidance and engage in the difficult but important direct engagement with values related to our health. If this is true of us, collectively, we might think it important to avoid nudging patients in either direction.

How do we know when we are in a context in which it is important for people to engage directly with values? These areas of decision will likely include “big personal decisions” that “address matters that comprise some of the primary bases of a meaningful or fulfilling life” (Tsai 2014), but might also include “space to make even trivial choices purely on the basis of the small, specific reasons that trivial options provide” (Shiffrin 2004). Returning to the example of the city’s health insurance plan, the decision about which health care plan to enroll in *might* be such a decision, but whether it deserves this sort of protection will depend on questions about how important it is for individuals to make decisions about health insurance and on how important it is to maintain relatively direct democratic control of the choice architecture surrounding such decisions. This is a question that we must resolve by examining the values of healthcare, and determining whether decisions about insurance are the sort that merit insulation from other considerations than those that have to do with health, not by examining how nudges, in general, impact decision-making by nudgees.

#### **IV. Conclusion**

The moral philosophers' objection to nudges claimed that nudges fail to adequately respect nudgees as choosers. I argued that it is implausible to interpret this objection to nudges as an objection to nudges reducing the number of options or choices that nudgees have, or an objection to nudges attempting to elicit a specific behavior other than through rational persuasion. A better interpretation of the objection to nudges is that nudges can make it more difficult for nudgees to democratically control the choice architecture in which they make decisions by remaining more hidden than other interventions, like mandates. But whether it is really objectionable to make choice architecture more hidden and more difficult to democratically control depends on how important the choices that the architecture shapes are to the people who make them. It is particularly in the context of choices where it is important for people to directly engage with a certain set of values that we might worry that nudges undermine democratic control of choice architecture, and this determination depends on substantive considerations about the situations in which people should have at least limited opportunities to engage with certain values directly. What is left of the moral philosophers' objection to nudges is not a claim that there is a *pro tanto* reason to avoid nudging, but that as a psychological matter, because nudges may be more hidden than other policy interventions, we might wish to pay special attention to the possibility that nudges can undermine the independence of decisions in choice contexts that ought to be insulated.

#### **References**

- Eyal, Nir. 2015. "Nudging and Benign Manipulation for Health."
- Hausman, Daniel M., and Brynn Welch. 2010. "Debate: To Nudge or Not to Nudge?" *Journal of Political Philosophy*. 18:123.

- Korobkin, Russell. 2009. "Libertarian Welfarism." *California Law Review*. 97:1651.
- Shiffrin, Seana Valentine. 2004. "Egalitarianism, Choice-Sensitivity, and Accommodation." in *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*, edited by R. Jay Wallace et al., 270–302. Clarendon Press. Oxford.
- Sunstein, Cass R. 2015. "Choosing Not to Choose."
- Tsai, George. 2014. "Rational Persuasion as Paternalism." *Philosophy and Public Affairs*. 42:78.
- Waldron, Jeremy. 2014. "It's All For Your Own Good." *New York Review of Books*. October 9.